Journal of Applied Statistics, Vol. 29, Nos. 1-4, 2002, 49-52

Discussion comments on: 'Occam's shadow: levels of analysis in evolutionary ecology—where to next?' by Cooch, Cam and Link

JAMES D. NICHOLS, *US Geological Survey, Patuxent Wildlife Research Center, MD, USA*

I thought this was an excellent paper filled with interesting ideas and suggestions worthy of serious consideration and contemplation. Here, I will focus discussion on two topics presented in this paper that pose investigative dilemmas and that are thus very relevant to the conduct of research in evolutionary ecology.

The first topic concerns aggregation and levels of analysis. Cooch *et al.* make a strong case that appropriate scales of aggregation and levels of analysis are important to developing reasonable inferences and should thus be determined by the nature of the question being addressed. I agree with Cooch *et al.* about the importance of decisions about levels of aggregation and analysis, and offer additional observations and thoughts on this subject.

I believe that some form of aggregation is necessary for the conduct of science. My reasoning is very simple. If we view an individual organism's fate or behaviour at any point in space and time as a unique event not capable of informing us about the likelihood of the event for other individuals or points in space and time, then generalization and prediction become impossible. The task of the biologist then involves simply recording and describing these unique events and possibly developing *a posteriori* stories to explain them (see Nichols, 2001). Although such descriptive work might be interesting, it is not consistent with most definitions of science. I do not claim, therefore, that nature cannot actually be a large collection of unique events; only that I prefer the job of science and will continue to hope that generalization and stochastic prediction are possible.

If we conclude that aggregation is necessary to conduct science, then we must still consider how to go about aggregating in ways that are likely to be useful, a topic considered by Cooch *et al.* It is well known that different ways of aggregating

Correspondence: US Geological Survey, Patuxent Wildlife Research Center, MD, USA.

ISSN 0266-4763 print; 1360-0532 online/02/010049-04
DOI: 10.1080/02664760120108449

or stratifying data can yield very different inferences, even reversing the rank order of the probabilities of interest (e.g. Simpson, 1951). An example of 'Simpson's paradox' was provided by Bickel *et al.* (1975), who investigated the sex-specificity of acceptance probabilities to graduate school at the University of California at Berkeley. The proportion of females who were denied admission was greater than that for men, leading to the natural charge of sex discrimination in admissions policy. However, when applications and admissions were analysed by department, women actually had higher probabilities of acceptance than males. The overall higher proportion of women denied admission had resulted from females differentially applying to departments with lower rates of acceptance.

Consider the estimation of a demographic rate parameter or fitness component such as survival probability for a population of interest. We can estimate a probability using all members of the population that begin the time interval of interest, but we would likely worry about the effects of heterogeneity on such estimation (e.g. see Carothers, 1973; Johnson *et al.*, 1986). Even if we were not worried about possible biases in estimates of 'average survival' resulting from heterogeneous probabilities among the aggregated individuals, we might be interested in individual characteristics that were the primary determinants of survival probability. Such interest would lead us to stratify the population by sex, age and other variables thought to influence survival. However, to some extent such stratification will be arbitrary and, because of Simpson's paradox, potentially misleading. One response to this arbitrariness and inferential difficulty is to stratify by a large number of potentially relevant covariates. However, the larger the number of strata, the fewer individuals in each stratum, and the more difficult it will be to estimate a stratum-specific survival probability. And we can never be certain that we have not overlooked an important covariate capable of reversing any inferences about the effects of covariates on survival. Of course, the logical extension of this approach of increasing stratification will yield a single individual in each stratum, with the corresponding estimation problem analogous to that of being asked to estimate the probability of heads from a single flip of a loaded coin. Cohen (1986) has referred to this dilemma as the 'uncertainty principle in demography' and provided a much more detailed discussion of the problem.

At first glance, the ability to model survival as a function of individual covariates (e.g. White & Burnham, 1999) may appear to rescue us from the consequences of the uncertainty principle. After all, use of individual covariates permits estimation of survival probability at the level of the individual animal. However, the reason why estimation in this situation differs from the single flip of the loaded coin involves a form of aggregation. By modelling survival probabilities as a function of a finite number of covariates, we are essentially aggregating over all other potential covariates and claiming that they are not relevant to the estimation of individual survival probability. Even in the case of the more flexible frailty (Shepard & Zeckhauser, 1980; Manton *et al.*, 1981) and random-effects (e.g. Link *et al.*, this issue; Burnham & White, this issue) modelling, we are still asserting that it is sensible to consider the individuals in the stratum of interest as being characterized by some distribution, thus borrowing information from other members of the collection (Link, 1999; Link *et al.*, this issue).

If, as suggested above, we are not able to escape the uncertainty principle, then how should we proceed in trying to draw inferences about survival probabilities and the variables that influence or determine them? In retrospective analyses, I simply believe that we must think harder about the selection of variables for use in stratification or covariate analyses. Rather than embarking on fishing expeditions with large numbers of variables, we should use all of our biological knowledge, insight and intuition in the selection of variables for study. This will not guarantee reasonable inferences, but is likely the best that we can do in retrospective analyses.

Having identified variables of potential importance using such analyses, the best way to draw strong inferences will involve manipulative experimentation. Nice examples of this approach are provided in this volume (e.g. Eicholz *et al.* this issue; Yoccoz *et al.* this issue).

The second topic concerns the interesting discussion in Cooch *et al.* of phenotypic plasticity and state-dependent decisions affecting fitness. In a discussion of possible costs of reproduction, for example, Cooch *et al.* noted that 'estimation is necessarily conditioned on the realized sequence of decisions; what may be more relevant in many cases are estimates of transition rates given a different sequence of decisions for that individual. However, clearly, this is not observable.' The argument here is that individuals that do reproduce at a specific time may differ in many important respects from individuals that do not reproduce, and that comparisons of survival rates for these two groups, for example, may be more relevant to these differences among individuals than to the event of reproduction. Individuals aggregated as reproducers and non-reproducers may differ in important variables other than the reproductive event, and these other variables may be very relevant to survival.

As was the case with the uncertainty principle in demography, I believe that there is no clear way to avoid this sort of problem in retrospective analyses of observational data. Once again, it seems that manipulative experimentation offers the only viable approach to strong inference. With such an approach, the intent would be to select individuals randomly and then impose treatments such as reproduction or non-reproduction, small or large clutch size, etc, to individuals in the two groups. Such an approach should yield strong inferences about the effects of the imposed treatments without risking the misleading inferences that sometimes result from unrecognized variables associated with observational or natural 'treatments'.

In summary, I believe that the discussion of levels of aggregation and analysis by Cooch *et al.* is very relevant to the conduct of science in the field of evolutionary ecology. The above comments simply suggest that for two classes of problem, strong inferences are not likely to result from retrospective analyses but will instead require manipulative experimentation.

REFERENCES

BICKEL, P. J., HAMMEL, E. A. & O'CONNELL, J. W. (1975) Sex bias in graduate admissions: data from Berkeley, *Science*, 187, pp. 398-404.

BURNHAM, K. P. & WHITE, G. C. (2002) Evaluation of some random effects methodology applicable to bird ringing studies. *In*: B. J. T. MORGAN & D. L. THOMSON (Eds), *Statistical Analysis of Data From Marked Bird Populations. Journal of Applied Statistics.*

CAROTHERS, A. D. (1973) The effects of unequal catchability on Jolly-Seber estimates, *Biometrics,* 29, pp. 79-100.

COHEN, J. E. (1986) An uncertainty principle in demography and the unisex issue, *American Statistician*, 40, pp. 32-39.

COOCH, E. G., CAM, E. & LINK, W. (2002) Occam's shadow: levels of analysis in evolutionary ecology—where to next? *In*: B. J. T. MORGAN & D. L. THOMSON (Eds), Statistical analysis of data from marked bird populations, *Journal of Applied Statistics*, this issue.

EICHOLZ, M. W., SEDINGER, J. S. & REXSTAD, E. (2002) Influence of nutrient reserves on annual survival of Canada Geese staging in interior Alaska, *Journal of Applied Statistics*, this issue.

JOHNSON, D. H., BURNHAM, K. P. & NICHOLS, J. D. (1986) The role of heterogeneity in animal population dynamics, *Proceedings of the International Biometrics Conference*, 13, 15 pp.

LINK, W. A. (1999) Modeling pattern in collections of parameters, *Journal of Wildlife Management*, 63, pp. 1017-1027.

LINK, W., COOCH, E. & CAM, E. (2002) Model-based individual fitness. *In*: B. J. T. MORGAN & D. L. THOMSON (Eds), *Statistical Analysis of Data from Marked Bird Populations*, *Journal of Applied Statistics*, this issue.

MANTON, K. G., STALLARD, E. & VAUPEL, J. W. (1981) Methods for comparing the mortality experience of heterogeneous populations, *Demography*, 18, pp. 389-410.

NICHOLS, J. D. (2001) On the use of models in the conduct of science and management of natural resources. *In*: T. SHENK & A. FRANKLIN (Eds), *Modeling in Natural Resource Management: Development, Interpretation and Application*. (Island Press), pp. 11-34.

SHEPARD, D. S. & ZECKHAUSER, R. J. (1980) Long-term effects of interventions to improve survival in mixed populations, *Journal of Chronic Diseases*, 33, pp. 413-433.

SIMPSON, E. H. (1951) The interpretation of interaction in contingency tables, *Journal of the Royal Statistical Society, Series B*, 13, pp. 238-241.

WHITE, G. C. & BURNHAM, K. P. (1999) Program MARK: survival rate estimation from both live and dead encounters, *Bird Study*, 46, pp. S120-139.

YOCCOZ, N., ERIKSTAD, K. E., BUSTNES, J., HANSSEN, S. & TVERAA, T. (2002) Reproductive costs in common eiders: the assessment of relationships between reproductive effort and future survival and reproduction based on experimental and observational studies. *In*: B. J. T. MORGAN & D. L. THOMSON (Eds), *Statistical Analysis of Data from Marked Bird Populations. Journal of Applied Statistics*, this issue.